FINITE GENERATION OF MAPPING CLASS GROUPS

SHINTARO FUSHIDA-HARDY

Abstract. Last week we saw an introduction to mapping class groups (of surfaces). This week we'll show that mapping class groups of surfaces are finitely generated, by studying its action on a certain "curve complex". If time permits, we'll study a specific generating set as well as its relations.

1. Overview of the talk

Given a smooth manifold with boundary $(X, \partial X)$, recall that its *mapping class group* is

$$\mathrm{MCG}(X, \partial X) = \mathrm{Diffeo}^+(X, \partial X)/\sim.$$

In words, this is the space of pointwise-boundary-fixing orientation preserving diffeomorphisms of X up to isotopy. Today we'll focus on the case where $X = \Sigma_g$, the closed genus g surface. In this case, it turns out the mapping class group is finitely generated.

Theorem 1.1 (Dehn-Lickorish theorem)**.** *For $g \geq 0$, the mapping class group $\mathrm{MCG}(\Sigma_g)$ is generated by finitely many Dehn twists about non-separating simple closed curves.*

Recall that a Dehn twist is a diffeomorphism of a surface supported in an annulus, which corresponds to "twisting" the surface by a full loop about the core of the annulus.

Remark. Dehn, who was one of the first people to study mapping class groups, was also the mathematician behind the *word problem*. This takes as input a group and a finite generating set for the group, and asks if there is an algorithm that takes as input a word in the generator and outputs whether or not the word is trivial in the group.

(1) There exist finitely generated groups with unsolvable word problem.
(2) For a given finitely generated group, solubility of its word problem is independent of the choice of generating set.

Theorem 1.2. *The word problem for mapping class groups is soluble.*

Proof. Let S be a finite collection of non-separating simple closed curves on Σ_g, such that Dehn twists about those curves generate the mapping class group. Let w be a word in these Dehn twists. The bigon criterion tells us exactly when curves are isotoped to minimise their intersections - the first step in the algorithm is to remove bigons via isotopy. Since the curves are in minimal position, we can apply the Alexander method. Specifically, consider the union of the curves as a graph on the surface. The Alexander method says that a mapping class $\varphi = w$ is trivial if and only if it fixes every edge and vertex of the graph! This is a finite check, so we've found an algorithm to solve the word problem. $\square$

So how will we show that mapping class groups are finitely generated? We'll use group actions, specifically the following pseudo-theorem:

Theorem 1.3 (Approximately a theorem)**.** *Suppose a group G acts on a connected space X. Let D be a subset of X with full orbit in X. Then G is generated by elements that only slightly shift D.*

Of course this needs to be formalised a lot more to make it true. The main idea we're trying to capture is that given the right conditions, every element of a group is a product of elements that incrementally translate D.

To show that mapping class groups are finitely generated, we'll construct a space on which the mapping class group acts, as well as a subspace D with full orbit. The mapping classes that "slightly shift D" turn out be finite. Specificaly, the space we'll construct is something called a curve complex.

2. Several examples of curve complexes

Definition 2.1. Let Σ be a surface. The *curve complex* $C(\Sigma)$ is the flag complex defined by the 1-skeleton given by the following data:

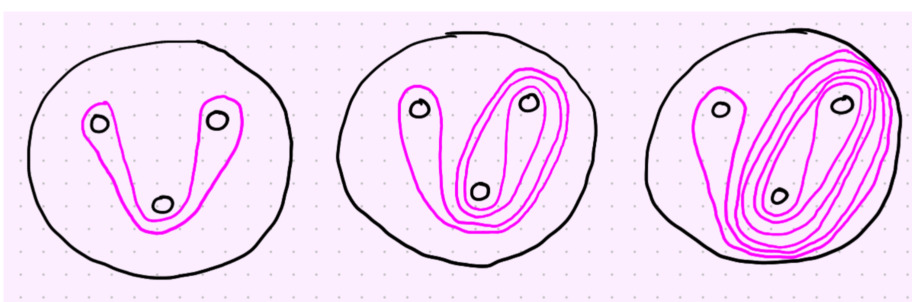
(1) There is a vertex for every isotopy class of essential simple closed curves in Σ.
(2) There is an edge between two vertices a and b if and only if $i(a, b) = 0$.

Example. Let Σ be a 2-sphere. Then every curve on Σ is isotopic to a point. Since no essential simple closed curves exist, there are no vertices. Therefore $C(\Sigma)$ is empty.

Example. Similarly, $C(\Sigma)$ is empty if Σ is a sphere with one, two, or three punctures.

Example. Suppose Σ is a sphere with four punctures. Then $C(\Sigma)$ has infinitely many vertices, but no edges.

To see why there are infinitely many vertices, we draw the following family of non-isotopic curves.



To see that the curve complex has no edges, consider two essential simple closed curves on Σ. Each one necessarily cuts Σ into two components with two punctures each. If these components consist of the same punctures, then the curves are isotopic. If they don't, then the curves must intersect.

Remark. One might say "hey, aren't those curves all isotopic?" referring to the claimed infinite family above. The reason these are truly distinct is that we can't drag the punctures around the disk: this would involve mapping classes that exchange punctures.

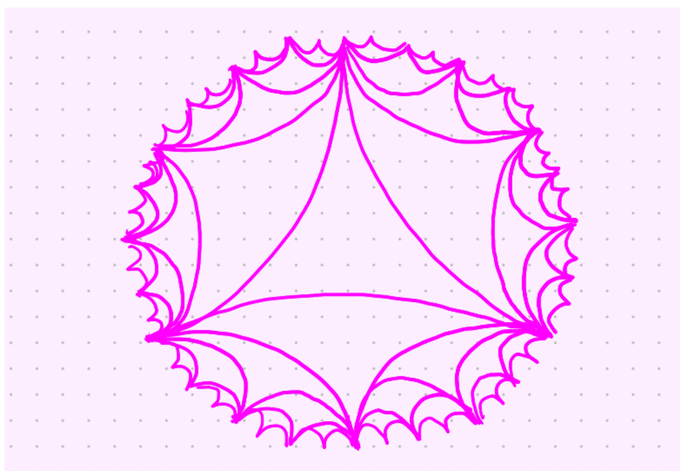
This means to study a four-times punctured sphere, we need a modified version of the curve complex. We introduce a few modifications here.

Definition 2.2. The complex $\check{C}(\Sigma)$ consists of the same vertices, but has edges whenever $i(a, b) = m$, where m is the minimum intersection number between any two non-isotopic essential simple closed curves on Σ.

Example. For sufficiently complicated surfaces, you can easily find two non-isotopic essential simple closed curves which don't intersect. This means $m = 0$, so $\check{C}(\Sigma) = C(\Sigma)$.

Example. For a four-times punctured sphere, we can draw two non-isotopic simple closed curves which intersect twice. It's not hard to see that one intersection is impossible. Therefore $m = 2$.

In fact, $\check{C}(\Sigma)$ is the *Farey complex*, which is a triangulation of hyperbolic space.



Definition 2.3. There's a subcomplex $N(\Sigma) \subset C(\Sigma)$ whose vertices are exactly *non-separating* essential simple closed curves on Σ. (That is, curves which don't cut Σ into two pieces when cutting along them.)

Definition 2.4. The actual curve complex used in the proof of finite generation of the mapping class group is $\widehat{N}(\Sigma)$, which consists of vertices isotopy classes of non-separating essential simple closed curves, and edges between a, b whenever $i(a, b) = 1$.

Recall that our goal was to show that the mapping class group is finitely generated by studying its action on the curve complex. This requires us to prove some properties of curve complexes.

Theorem 2.5. *When $3g+n \geq 5$, then $C(\Sigma_{g,n})$ is connected. (Here $\Sigma_{g,n}$ means the surface with genus g and n punctures.)*

Proof. I won't give a proof but I'll sketch the idea. Suppose a and b are curves on $\Sigma_{g,n}$, i.e. vertices of the curve complex. If $i(a, b) = 0$ then there's an edge from a to b so we're done. In general, we want to find a collecton of curves $a = c_0, c_1, \ldots, c_n, c_{n+1} = b$ so that $i(c_m, c_{m+1}) = 0$. We proceed inductively:

(1) In the case $i(a, b) = 1$, we find a curve c so that $i(a, c) = i(c, b) = 0$.
(2) In the case $i(a, b) = n \geq 2$, we find a curve c so that $i(a, c), i(c, b) < n$.

I won't provide all the details, but it goes something like this:

Base case. Consider a neighbourhood of the union of a and b. These intersect exactly once, so it looks a bit like "gluing two annuli together". Topologically, this is homeomorphic to a torus with a disk removed. In particular, the boundary of this surface is a curve on $\Sigma_{g,n}$ which doesn't intersect a or b. It remains to check that this curve is actually an *essential* simple closed curve; i.e. that it isn't null-homotopic or homotopic to a point. If it were, then the surface is necessarily $\Sigma_{1,0}$ or $\Sigma_{1,1}$, both of which violate the premise that $3g + n \geq 5$.

Inductive step. We won't provide any details - the idea is similar to above. Essentially, the curves a and b intersect at least twice. We can do surgery near these two intersections to find a curve c which decreases the number of intersections. □

Theorem 2.6. *$N(\Sigma_{g,n})$ is connected if $g \geq 2$.*

Theorem 2.7. *$\widehat{N}(\Sigma_{g,n})$ is connected if $g \geq 2$.*

The proof has been omitted.

3. Quasi-stabiliser generation theorems

These theorems don't actually have names! I just decided to call them quasi-stabiliser generation theorems. I'm now going to give formal statements of some theorems.

Theorem 3.1. *Let G act on a connected topological space X. Suppose D is an open* translation domain*. Then* QStab(D) *generates G.*

Here a *translation domain* is a subset $D \subset X$ such that $GD = X$. This is actually a general case of a *fundamental domain* which is a phrase we hear often in geometry. For example, the famous fundamental domain of hyperbolic space is a strange triangle in the upper half plane model. (Precisely, a fundmental domain is a translation domain with the added property that gD only ever meets D along their boundary for g non-trivial.)

Next, QStab(D) is what I've called the *quasi-stabiliser*. It's defined to be

$$\mathrm{QStab}(D) \coloneqq \{g \in G : gD \cap D \neq \varnothing\}.$$

Recall that the *stabiliser* of a point $x \in X$ under a group action by G is

$$\mathrm{Stab}(x) \coloneqq \{g \in G : gx = x\}.$$

We see that the quasi-stabiliser is something that "stabilises points in the domain, up to the domain".

Proof. We consider the space $gD \cap \langle \mathrm{QStab}(D) \rangle D$ for some $g \in G$. Suppose it's non-empty. (Such a g exists - the identity element works.) It follows that there exists some s in $\langle \mathrm{QStab}(D) \rangle$ for which $gD \cap sD \neq \varnothing$, so $s^{-1}gD \cap D \neq \varnothing$. This is exactly the definition of $s^{-1}g$ being in $\mathrm{QStab}(D)$! Thus

$$g \in s\,\mathrm{QStab}(D) \subset \langle \mathrm{QStab}(D) \rangle.$$

It follows that $(G - \langle \mathrm{QStab}(D) \rangle)D$ is disjoint from $\langle \mathrm{QStab}(D) \rangle D$. (This is exactly what we've shown - we started with an arbitrary g and showed that if gD meets $\langle \mathrm{QStab}(D) \rangle D$, then g is also in $\langle \mathrm{QStab}(D) \rangle$. Thus, it *isn't* in $G - \langle \mathrm{QStab} \rangle$. Since D is open, each of the translations is open. Moreover, their union is X. From point-set topology, if two open sets union to the whole space, then they're distinct connected components! Because X is connected, and $\langle \mathrm{QStab} \rangle D$ is non-empty, $(G - \langle \mathrm{QStab}(D) \rangle)D$ is empty. It follows that $G - \langle \mathrm{QStab}(D) \rangle$ is empty, i.e. $\langle \mathrm{QStab}(D) \rangle$ is all of G. $\square$

Next we introduce another version of the theorem which we're actually going to use.

Theorem 3.2. *Let X be a connected simplicial complex, and G a group that acts on X by simplicial automorphisms. (This basically means vertices map to vertices, edges to edges, etc.) Suppose D is a subcomplex of X which is a translation domain. Then* $\mathrm{QStab}(D)$ *generates G.*

This is genuinely distinct from the previous version of the theorem, because subcomplexes of simplicial complexes aren't open.

The proof will be omitted, but it's very different from the previous proof! It can be proven by induction, by considering a sequence of vertices between v and gv (for any g), and showing that each sucessive intermediate vertex differs by an element of $\mathrm{QStab}(D)$ from the previous one.

Actually the version we're really going to use is some sort of a special case: the statement is slightly stronger, but the proof (which is also omitted) is pretty much identical to the previous proof.

Theorem 3.3. *Let X be a connected simplicial complex, and G a group that acts by simplicial automorphisms. Suppose G acts transitively on vertices and edges. Let u and v be adjacent vertices in X and $h \in G$ such that $hu = v$. Then*

$$\{h\} \cup \mathrm{Stab}(v)$$

generates G.

4. Finite generation of mapping class groups

Finally we give a proof outline to show that mapping class groups are finitely generated.

Theorem 4.1. $\mathrm{MCG}(\Sigma_g)$ *is finitely generated (by Dehn twists about non-separating simple closed curves).*

Proof.
(1) If g is less than 2, then the surface is a sphere or a torus. In the case of the torus, the mapping class group $\mathrm{SL}(2, \mathbb{Z})$ is generated by Dehn twists about the meridian and longitude. The sphere is trivial.

(2) For $g \geq 2$, we have that $\widehat{N}(\Sigma_g)$ is connected. The mapping class group acts by simplicial automorphisms (since it sends isotopy classes of non-separating curves to other isotopy classes of non-separating curves). By the last quasi-stabiliser generation theorem, we have that

$$\{h\} \cup \mathrm{Stab}(v)$$

generates $\mathrm{MCG}(\Sigma_g)$, where v is any vertex of $\widehat{N}(\Sigma_g)$, and h is an element of G sending v to an adjacent vertex.

(3) We proceed by induction: suppose $\widehat{N}(\Sigma_{g-1})$ is generated by Dehn twists about finitely many non-separating essential closed curves. Choose any vertex v in $\widehat{N}(\Sigma_g)$, i.e. any non-separating essential simple closed curve in Σ_g. Choose u to be an adjacent vertex, i.e. a curve intersecting v exactly once. By a property of Dehn twists,

$$hu := T_u T_v(u) = v,$$

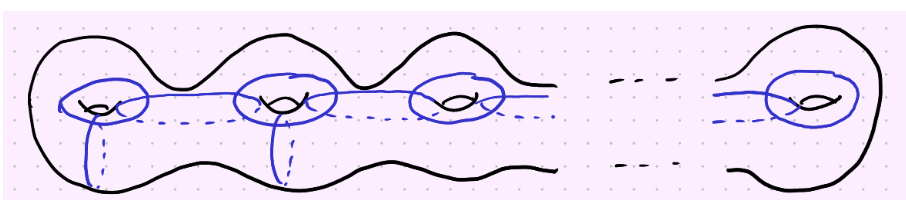
so $\mathrm{MCG}(\Sigma_g)$ is generated by $T_u T_v$ and the stabiliser of v.

(4) The stabiliser of v is generated by the Dehn twist about v, and the mapping class group of the surface obtained by cutting along v. The latter surface has genus $g-1$, so the inductive hypothesis applies!

Technically we needed to induct over both n and g. In the last step, cutting along v introduces a new boundary component (puncture). A full proof first inducts over the number of punctures, and then the genus. □

5. The Humphries generators

The Humphries generators are an explicit set of non-separating essential simple closed curves whose Dehn twists generate the mapping class group.



Theorem 5.1. *The Humphries generators (of which there are $2g+1$) generate* $\mathrm{MCG}(\Sigma_g)$*. In fact, any generating set consisting of Dehn twists must have at least $2g+1$ generators.*

We now give a proof outline that $2g+1$ generators is minimal.

Proof. (1) The *symplectic group* $\mathrm{Sp}(2g, \mathbb{Z})$ consists of $2g \times 2g$ matrices preserving a non-degeneration skew-symmetric bilinear form.

(2) Any mapping class induces an isomorphism on homology; i.e. an element of

$$\mathrm{Aut}(H_1(\Sigma_g)) = \mathrm{Aut}(\mathbb{Z}^{2g}) = \mathrm{GL}(2g, \mathbb{Z}).$$

(3) Mapping classes preserve the *algebraic intersection number* $\hat{i}(a, b)$. This is a non-degenerate skew-symmetric bilinear form, so we actually have a map

$$\mathrm{MCG}(\Sigma_g) \to \mathrm{Sp}(2g, \mathbb{Z}).$$

(4) This is actually a surjection. Overall we can construct a surjection

$$\mathrm{MCG}(\Sigma_g) \to \mathrm{Sp}(2g, \mathbb{Z}/2\mathbb{Z}).$$

In this setting there's an elementary proof that at least $2g + 1$ generators are required.

$\square$